

AMENDMENTS

Please amend the claims as follows:

1. (Claims 1-9 are cancelled.)

10. (Currently amended) An unnecessary word determination method in a document automatic classification system, comprising the steps of:

generating a word list for each of at least two categories by extracting words from a learning document set, the word list containing information on a frequency of appearance of each extracted word within each category and a part of speech corresponding to each word;

filtering the word list to eliminate the words that are identified as a particular part of speech;

determining an unnecessary word for a first category on the basis of the number of occurrences of the word within at least one other category wherein a word is determined to be unnecessary in the first category in response to the word having a greater number of occurrences than a given standard in ~~the at least~~ one other category, a distinct given standard defined for each category, the given standard defined irrespective of the number of occurrences of the word in the first category, the given standard defined as the product of the number of documents in a corresponding category and a predefined threshold value;

generating a document classification catalog by eliminating words determined to be unnecessary words from each of the word lists, the document classification catalog comprised of a plurality of vector spaces wherein each vector space represents at least one category, each vector space comprising the number of occurrences of the remaining words in the word list of a corresponding category.

receiving a target classification document and generating a document vector for the target classification document, wherein a distance is defined between the document vector and each of the plurality of vector spaces such that the distance indicates a degree of similarity between the target classification document and a category represented by the vector spaces;

identifying the category corresponding to the vector space with the least distance between the

vector space and the document vector; and
assigning the target classification document to that category.

11. (Canceled).

12. (Canceled).

13. (Canceled).

14. (Currently amended) An unnecessary word determination method in a document automatic classification system, comprising the steps of:

acquiring information on words from a document set, classifying the words according to category, and storing the words in a storage device;

filtering the words to eliminate the words that are identified as a particular part of speech;

recognizing the number of occurrences within at least one other category of a word belonging to a given category on the basis of the acquired information;

determining an unnecessary word for a first category on the basis of the number of occurrences of the word within at least one other category wherein a word is determined to be unnecessary in the first category in response to the word having a greater number of occurrences than a given standard in ~~the at least one other category~~, a distinct given standard defined for each category, ~~the given standard comprised of a predetermined threshold scaled by the number of documents in the at least one other category and~~ defined irrespective of the number of occurrences of the word in the first category, the given standard defined as the product of the number of documents in a corresponding category and a predefined threshold value; and

generating a document classification catalog by eliminating words determined to be unnecessary words.

15. (Previously presented) The method according to Claim 14, further comprising storing said classification catalog into the storage device.

16. (Previously presented) The method according to Claim 15, further comprising the step of performing classification processing for classification target documents by using the classification catalog stored in said storage device.